Office of Environmental Health Hazard Assessment

Winston H. Hickox
*Secretary for
Environmental
Protection*

Gray Davis
Governor

April 14, 1999

Mr. Daniel S. Greenbaum, President
Health Effects Institute
955 Massachusetts Avenue
Cambridge, Massachusetts 02139

Dear Mr. Greenbaum:

This letter is to bring to your attention our concerns about the Health Effects Institute's (HEI) diesel epidemiology panel's draft report presented March 7-9, 1999 at Stone Mountain, Georgia, during the <u>Diesel Workshop: Building a Research Strategy to Improve Risk Assessment</u>.

Our concerns, both technical and procedural, are as follows:

♦ Choice of overly restrictive criteria for selection of epidemiology studies to be used for quantitative risk assessments for diesel exhaust;

♦ Public health protection and Office of Environmental Health Hazard Assessment's (OEHHA) health assessment for diesel exhaust particulate inadequately considered;

♦ Peer review comments on the draft report inadequately addressed; and

♦ Important implications of the HEI, diesel epidemiology panel's findings.

Criteria for Selection of Epidemiology Studies

We believe that the panel's report and findings presented at the conference were based on overly restrictive criteria for using epidemiological studies for purposes of quantitative risk assessments. The criteria are set at near perfection and are so restrictive as to preclude virtually any existing epidemiological study (not just those examining effects of diesel exhaust) from being used for quantitative risk assessment. If the standard is set too high for use of epidemiological studies in risk assessment, agencies charged with protecting public health will have to rely exclusively on animal studies resulting in greater uncertainty. We recognize that even with

inherent uncertainties in epidemiology studies, quantitative risk assessments can be developed from such studies.

Public Health Protection and OEHHA Risk Assessment

Considering how widespread exposure to diesel exhaust is, scientific groups and government agencies have the responsibility to try to assess the dimensions of risks to public health and, if appropriate, reduce exposures. Although speakers at the workshop, including yourself, agreed that the experimental and epidemiological data support the designation of diesel exhaust as a lung carcinogen, the study criteria, in being overly restrictive, did not account for the importance of public health protection. Without this consideration the effort becomes an academic exercise, removed from the practical implications and health implications of the panel's findings. This can undermine the ability of regulatory agencies to undertake efforts to protect public health by reducing risk.

The Part B: Health Assessment for Diesel Exhaust Particulate prepared by OEHHA and which used epidemiology studies for the quantitative risk assessment was approved unanimously by the California Air Resources Board after a nine year, iterative, open process. It involved many comments and written responses, several workshops (including the scientific workshop coordinated in part by HEI staff), an extended review by our independent Scientific Review Panel, and several special legislative hearings. I believe that the HEI panel in preparing their findings did not seriously consider this body of work and the substantial amount of public and scientific input.

Peer Review

Preparation of peer review comments on the draft report consumed a considerable amount of staff time and, considering the potential consequences of the findings, I believe HEI was remiss in not allowing more time for substantive rebuttal of the panel's findings at the meeting. I am enclosing a copy of the technical comments sent to HEI by our staff. These comments present specific rebuttals to several of the panel's findings, which we urge HEI and the panel to consider in drafting the final report.

Implications of Findings

Together with a previous HEI report finding against using animal studies for quantitative risk assessment, the findings of this panel appear to imply a conclusion that a risk assessment for diesel exhaust exposure cannot be conducted. Many individuals seem to interpret this conclusion to mean that there is no risk from diesel exhaust exposure. It is imperative that the HEI panel clarifies its position regarding human risk and the protection of public health from diesel exhaust exposure.
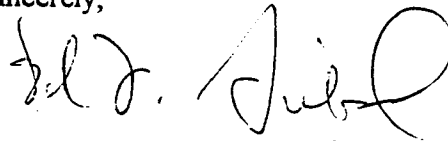
Many of the HEI conference attendees, including ourselves and a wider audience around the world, have an enormous stake in the panel's conclusions. Some organizations will misconstrue the preliminary findings of the HEI panel. One example is provided in the enclosed article from the newsletter of the California Trucking Association: "The elite panel discredited the work done by California as scientifically inadequate and unanimously determined that the cancer slope found by the California's Science Review Panel was erroneous." This is untrue. However, it does point out the importance of HEI and the epidemiology panel ensuring that all scientific elements of the diesel exhaust debate are methodically and seriously considered.

In closing, I want to urge you and the panel to reconsider the overly stringent criteria for diesel exhaust epidemiology studies, consider public health implications in adopting such criteria, and consider the comments of all peer reviewers.

If you would like to discuss these issues, please call me at (916) 322-6325.

Sincerely,

Joan E. Denton, Ph.D.
Director

Enclosures

cc: Alan C. Lloyd, Ph.D., Chairman (w/enclosures)
Air Resources Board
P.O. Box 2815
Sacramento, California 95812

COMMENTS ON THE PRESENTATION OF
THE HEI DIESEL EPIDEMIOLOGY EXPERT PANEL.

These comments supplement my February 18 review of the panel's draft report. They focus on aspects of the panel's presentation on March 8 and include some of my remarks from the floor of the workshop during SESSION IV: What do published studies tell us about exposure-response?

Panel Presentation

John Bailar presented the panel's conclusion that the Garshick et al. (1987, 1988) railroad worker data are unsuitable for quantitative risk assessment (QRA). The principal rationales underlying the panel's decision were: (1) inadequate data, including the retrospective exposure assessment, and (2) calculation of a non-monotonic exposure-response relationship, possibly due to uncontrolled confounding (e.g. smoking), exposure misclassification, the healthy-worker-survivor effect or incomplete case ascertainment.

1) Marie Swanson presented a framework for evaluating epidemiological studies for QRA. She provided stringent criteria for studies to be suitable for risk assessment. These criteria included provisions for measures of smoking using a case-control study nested within the cohort under study. The criteria also included a meticulous, thorough exposure assessment.

2) David Hoel presented results of an exploratory analysis of the railroad cohort data. He found overall that train riders had significantly elevated risks and that for clerks alone, train riders alone and shop workers alone, the risk of lung cancer decreased with duration of exposure. Analysis included controls for attained age and calendar year. When a dichotomous variable (GRP) classifying workers as exposed or unexposed according to job was included the analysis for the ramp exposure pattern with background correction, the slope of risk with exposure went from significantly positive to insignificantly negative.

Response

My response to Dr. Bailar's presentation of the finding of unsuitability of the railroad worker cohort for quantitative risk assessment is to point out how neither the presentation given by Dr. Swanson nor the presentation given by Dr. Hoel supports that finding in the actual world of risk assessment.

1) The panel's criteria for epidemiology studies seem to set near perfection for a goal, but there are no nearly perfect epidemiology studies. So criteria are needed for the degree to which that goal needs to be met in order for a risk assessment to be useful.

   ♦ Different purposes for QRAs. There are different purposes for QRAs, and the degree of implementation of the goal needs to depend on the purpose of the QRA

as well as the quality of the available evidence. Hertz-Picciotto (1995) designated 3 categories of studies to meet QRA needs. Her first and most stringent category is for those studies that are suitable for extrapolation. Within that category, I propose subcategories based on the goal or context of the risk assessment needs. The subcategories would be closely related to the level of uncertainty of the supporting studies as well as the public health impact of exposure: (1) immediate danger to health, (2) moving toward a goal of health protection, and (3) early warning about the potential magnitude of a health risk. It appears that the diesel exhaust QRA most closely matches the second category.

♦ Need for following uniform guidelines. Regulatory risk assessments follow guidelines that afford a degree of uniformity among assessments of various substances. If the stringency placed on suitability of health studies for diesel exhaust is too much greater than that for other substances, then the regulation of diesel exhaust as an air pollutant becomes skewed in comparison to other air pollutants. The discussion and debate that occurs in formulating regulatory guidelines is an appropriate forum for discussing the stringency issue.

♦ Need for action. In the regulatory sphere, action cannot always wait until the nearly perfect study appears. The practice of risk assessment recognizes the principle of health protection in the light of uncertainty. The question is, "How much uncertainty?" The answer can make use of criteria such as those of Hertz-Picciotto.

♦ Bracketing exposure. The panel's criteria for exposure assessment are far beyond what is typically found in occupational epidemiology. The QRA by Cal/EPA reasonably bracketed the possible exposure of workers to diesel exhaust particles. Based on national average estimates of Woskie et al. for the early 1980s, the report provided calculations of risk based on two different patterns of exposure concentration for the previous 25 years, the ramp and the roof. Concentrations could not reasonably have been less than those of the ramp nor much greater than those of the roof. Thus, the argument against using these data because of lack of concurrent exposure information is moot.

2) In Dr. Hoel's presentation, the exploratory analysis sets forth a common basis for the analyses of Garshick, Crump and Dawson. The analysis used empirical models with exposure characterized either as duration or as a ramp pattern of concentration. That analysis explored use of the dichotomous exposed-unexposed variable (GRP) but did not explore the use of the GRP variable for the most realistic exposure pattern, the roof, and for the most appropriate models, multistage.

♦ The panel's use of the GRP variable. The use of the GRP variable by the panel was an appropriate exploration to report because it significantly improved the fit of the model to the data. Also, GRP allows to some degree an exploration of the exposure-response relationship without including the unexposed group, in case the

clerks and train riders are not homogeneous. However, if the worker population is homogeneous, then the inclusion of the GRP variable has some effect of double-counting the exposure variation. Thus the use of the GRP variable must be considered dubious, especially from the point of view of health protection. Because the use of the GRP variable apparently had a dramatic impact on the slope result for the model and exposure pattern considered by the panel, it is imperative that the panel evaluate the importance, relevance and implication of using the GRP variable in the analysis.

♦ <u>Use of the GRP variable for roof exposures.</u> The exploratory analysis using the dichotomous GRP variable was too narrow to draw sweeping conclusions without including other exposure patterns and models. Preliminary calculations we conducted using this variable show that its addition to the analysis of the roof pattern of concentration in an empirical model or in multistage models used in the Cal/EPA report resulted in no significant improvement of model fit. Therefore, the GRP variable should not be included in analyses using the roof pattern, which is the most realistic pattern. Without the GRP variable, these analyses all find very significant positive slopes. These results, both with and without GRP, are documented in the table below. The basic analyses without the added variable are found in the Cal/EPA risk assessment. For the convenience of the panel I sent by overnight mail our manuscript detailing the multistage portion of the analysis without the added variable. It would seem incumbent on the panel to explore the full range of plausible exposure patterns and models before coming to a conclusion about the suitability of the data for QRA.

♦ <u>Implications of the panel's approach.</u> One must also look at the implications that the more general results of Dr. Hoel's calculations may have for public health. Dr. Hoel confirmed an elevation in risk for the train riders, who were exposed to diesel exhaust. This result is consistent with the great body of literature suggesting that diesel exhaust exposure increases the risk of lung cancer in humans. The level of evidence for diesel exhaust exceeds that of many other substances identified as carcinogens and of concern to public health officials. In California, we were mandated to estimate the risk from diesel exhaust exposure in order to protect public health. The HEI panel needs to consider the actual context of diesel exhaust exposure. The present choices for evaluating risk from diesel exhaust are: (1) use of the animal data, (2) use of the human data, (3) not to conduct a risk assessment.

♦ Probability calculation. The panel might consider thinking about choices in quantitative risk assessment in terms of quantitative probability calculations. The values for the probabilities may be difficult to document and use in an ultimate presentation, but the arithmetic is simple and the thought process is instructive. Considering only two outcomes for simplicity, the basic formula is

$\mathcal{E}$ (risk) = $\mathcal{E}$ (risk | A) P(A) + $\mathcal{E}$ (risk | B) P(B), where

P(A) + P(B) = 1, and

$\mathcal{E}$ (risk) is the expected value of risk of lung cancer due to diesel exhaust,
$\mathcal{E}$ (risk | X) is the expected value of risk, given the slope calculation X,
P(X) is the probability that the slope calculation X is correct.

In actuality many other outcomes, for example those from a range of exposure scenarios, would be included in such a formulation. But for present purposes, consider applying this formula to the results of the calculations described in the table below and in the accompanying remarks. In that context let A be the outcome for use of the GRP category, and let B be the outcome without it.

For the general empirical model using the ramp pattern of exposure, let $\mathcal{E}$ (risk | A) =0 and $\mathcal{E}$ (risk | B) = 2.7 x 10^{-4}, from the table below (essentially equivalent to 3.2 x 10^{-4} of the panel's calculation, as described in the Methods Section of the Remarks for the Table). The probability P(B) that the calculation without the GRP variable is correct choice may be small, but it is not prudent simply to put value of P(B) at zero. If it is not put at zero, then the formula shows that the expectation of the risk is not zero. Reporting otherwise in a QRA would not present a fully scientific conclusion.

For the general empirical model using the roof pattern of exposure, let $\mathcal{E}$ (risk | A) = 0 and $\mathcal{E}$ (risk | B) = 1.1 x 10^{-4} from the table below. P(A) should now be put at a small though non-zero value. Accordingly P(B) would be near one. Then the risk is essentially as given in the Cal/EPA report. A similar result applies for the multistage model with roof exposure pattern.

Conclusion

In response to the title of this session, the most appropriate analyses and selections of the very imperfect data from the Garshick et al. (1988) cohort study can tell us a lot about the exposure-response. They can give us a reasonable indication of the magnitude of likely linear increases of risk with exposure.

Remarks on the attached table

Methods
The analysis uses multiplicative models without interactions. Except for the dichotomous variable GRP indicating whether a worker was a train rider (exposed) or a clerk-signalman (unexposed), the analysis and variables are as in Cal/EPA (1998) Appendix III, Part B, Appendix D. The other covariates are
Byr = birth year, equivalent to age in 1959,
Cyr = calendar year.

Age is taken into account in two different ways according to type of model. (1) In the general empirical model, lung cancer rates are referred multiplicatively to standard US values. (2) In the multistage model, a power of age is used, according to the Armitage-Doll approximation. Note that even the general model with a ramp exposure pattern is not exactly comparable to that used by Dr. Hoel's analysis, which used attained age as a covariate rather than referring rates to US standard values.

For calibration purposes, note that the risk slope for the multiplicative covariate combination, Byr and Cyr, without the GRP covariate, is consistent with slopes presented by Dr. Hoel for the highly adjusted cases without the GRP covariate. Thus, converting the values in the table here, obtained for an exposure concentration of 50 $\mu g/m^3$, to the equivalent value obtained in his analysis using a concentration of 43 $\mu g/m^3$, gives $2.7 \times 10^{-4} \times 50 / 43 = 3.1 \times 10^{-4}$, which is very close to the values of 3.2×10^{-4} to 3.3×10^{-4} for the nearest corresponding cases of his table.

Results
Using the ramp pattern in the general empirical model, the nested likelihood criteria retain the GRP variable as significantly contributing, along with both the other covariates, to the fit of the model to the data. This choice of covariates produces a non-significant slope of risk with exposure (boxed entry in the first row of the first column of values in the table below).

Using the roof pattern in the general model and in the two forms of the multistage model, these criteria show that the GRP variable does not contribute significantly of the fit of the models to the data. The other two covariates do contribute significantly. The slopes that result from using these covariates are significantly positive (boxed entries in the second through the fourth row of the fourth column of values in the table below).

Although the slopes for all the possible combinations of these covariates are shown in the table below, the boxed entries are the ones significantly preferred using goodness of fit criteria.

5

April 8, 1999

MLE SLOPES FOR RISK ON CUMULATIVE EXPOSURE
FOR DIFFERENT MODELS, COVARIATES AND EXPOSURE PATTERNS[abc]

Exposure Pattern	Model	All 3 covariates	Cyr+ GRP	Byr+ GRP	Byr+Cyr	Byr	Cyr	GRP
ramp,	general empirical	(slope NS) All 3 covariates significant	(slope NS)	(slope NS)	2.7E-04	2.9E-04	2.1E-04	(slope NS)
roof,	general empirical	(slope NS) GRP NS; other 2 significant	(slope NS)	(slope NS)	1.1E-04	1.2E-04	8.8E-05	(slope NS)
roof,	6/7 multistage	(slope NS) GRP NS; other 2 significant	(slope NS)	2.4E-03	1.5E-03	1.7E-03	1.7E-03	3.6E-03
roof,	7/7 multistage	(slope NS) GRP NS; other 2 significant	(slope NS)	2.4E-03	1.5E-03	1.9E-03	1.7E-03	3.9E-03

a. NS: not statistically significant. MLE: maximum likelihood estimate.
b. These are raw slopes. Numerical values are shown only for statistically significant slopes. The scale for the general empirical models in the first two rows differs from the scale for the multi-stage models in the last two rows. See Cal/EPA (1998) Appendix III, Part B, Appendix D.

S. Dawson

April 8, 1999

c. Boxes highlight the value of slope selected using nested maximum likelihood criteria. Rationale for selection is in the first slope column, considering all 3 covariates.

California Diesel Toxicity Challenged By Nationally Recognized Scientists

Railroad study found Inadequate for Risk Assessment

By Stephanie Williams
CTA Director of Environmental Services

The Health Effects Institute (HEI) is a congressionally mandated research organization that evaluates the science on air pollution for Congress. It is funded jointly by federal EPA and the international engine manufacturers. HEI held a Diesel Health Conference in Georgia March 7-9, 1999. This first time diesel health conference included a panel review of the underlying study that led California to list particulate emissions from diesel exhaust a toxic air contaminant.

The HEI expert panel that evaluated the railroad workers study presented their preliminary findings on the use of this study. The panel consisted of university researchers who had not previously worked with the diesel literature so that the "true believers" were not evaluating the science. The elite panel discredited the work done by California as scientifically inadequate and unanimously determined that the cancer slope found by California's Science Review Panel was erroneous. The independent HEI panel did give California a chance to respond in the meeting. Stan Dawson, staff statistician for the Office of Environmental Health Hazard Assessment (OEHHA), disagreed with the panel but had no meaningful scientific arguments to add to the panel decision.

The panel went further than challenging the science behind California's unit risk number (estimated deaths per million due to diesel exposure), the number that gave Proposition 65 suits their juice from the beginning against trucking fleets. The panel challenged that the railroad workers cohort study did not adequately provide controls for smoking and found the study inad-

The panel determined that estimating exposure data collected decades after the exposure occurred was not dependable when developing a relationship between diesel exposure and lung cancer.

equate for even hazard identification. The author, Dr. Garshick, also addressed the panel and spoke to their concerns regarding control for smoking. The panel found that smoking has such an enormous effect on lung cancer that detail information should have been acquired in the initial study, not in a later study.

The critical failing of the railroad study was the lack of concurrent exposure data. The panel determined that estimating exposure data collected decades after the

exposure occurred was not dependable when developing a relationship between diesel exposure and lung cancer. It appeared the health advocates were accepting that diesel may not cause cancer and were shifting towards non-cancer endpoints (asthma, lung irritation) and other studies that would prove their point.

The abandonment of the railroad workers study has lead federal EPA and other government agencies to embrace a more recently released Teamsters Union study on truck drivers authored by Kyle Steenland and Dennis Zaebst. Steenland's finding are worse than those found in the railroad workers study. The HEI panel also addressed the Teamsters study, pointing out smoking as a confounding variable and problems quantifying exposure.

The most important task for the trucking industry is to quantify the fleet during the period that the Teamsters study spanned. It could be that this study measured Teamster drivers that were exposed to gasoline exhaust. The first diesel engines were introduced in the late 1950's and many companies did not completely convert their fleets until the late seventies. The trucking industry has the challenge of recreating the national fleet during the forties and fifties and we intend to get started right away!